

EFFECT OF RESPONSE PRACTICE VARIABLES ON LEARNING SPELLING AND SIGHT VOCABULARY

ANTHONY J. CUVO, KRISTIN M. ASHLEY, KIMBERLY J. MARSO,
BINGJU LUCY ZHANG, AND TROY A. FRY

SOUTHERN ILLINOIS UNIVERSITY AT CARBONDALE

Four experiments were conducted to examine variables associated with response practice as an instructional technique for individuals with intellectual disabilities. In Experiment 1, the effect of the cover component in the "cover write" method was evaluated, as were the comparative effects of written versus oral practice of spelling words by rehabilitation clients. The results showed that the cover procedure generally did not enhance performance over and above that produced by practice alone, and written practice generally was not superior to oral practice. Experiment 2 demonstrated that less response practice (i.e., five times) was as effective as more practice (i.e., 10 and 15 times) for teaching spelling to adolescents with developmental disabilities. Experiments 3 and 4 also showed that even less response practice (i.e., one time) was as effective as more practice (five times), and irrelevant practice following errors was as effective as relevant practice for teaching spelling and sight vocabulary to adolescents with behavior disorders and developmental disabilities, respectively. The findings suggest that a parsimonious procedure of limited response practice and positive reinforcement may be effective for the tasks and populations studied.

DESCRIPTORS: response practice, mental retardation, spelling, sight vocabulary, developmental disabilities

In a review of the research on spelling instruction, the conclusion was drawn that the traditional classroom approach to teaching spelling vocabulary (i.e., assigning students to learn to spell 10 to 20 unrelated words at the beginning of the week and testing them at the end of the week, either with or without some form of group instruction) has not been successful (Heron, Okyere, & Miller, 1991). Specialized discrete-trial approaches derived from behavioral principles have been characterized as offering greater instructional efficiency and effectiveness.

Several error-correction procedures, termed variously as *overcorrection*, *positive practice*, and *directed rehearsal*, have been investigated in the

context of discrete-trial academic instruction of students with disabilities. Research has shown the efficacy of repeated practice in various forms on the acquisition of academic tasks, including spelling (Axelrod, Kramer, Appleton, Rockett, & Hamlet, 1984; Baker, 1992; Dalrymple & Feldman, 1992; Foxx & Jones, 1978; Mabee, 1988; Matson, Esveldt-Dawson, & Kazdin, 1982; Ollendick, Matson, Esveldt-Dawson, & Shapiro, 1980; Singh, Farquhar, & Hewett, 1991; Stewart & Singh, 1986) and sight vocabulary (Baker, 1992; Singh, 1987; Singh & Singh, 1986, 1988; Singh, Singh, & Winton, 1984) by students with disabilities. The purpose of the present series of experiments was to examine several variables that may affect practice as an instructional procedure in the discrete-trial instruction of children and adults with disabilities. Variables investigated included intermittently covering a written model of a word to be learned (Experiment 1), written and oral practice (Experiment 1), frequency of practice (Experiment 2), and task relevance of practice (Experiments 3 and 4).

We thank Sue Plumlee, Leslie Villani, and Guy Weber for their cooperation during the research, Edward Favorite, Thomas Gaffaney, and Donna Walden for serving as secondary scorers, and Paula Davis, Brandon Greene, Roger Poppen, and Harris Rubin for comments on an earlier draft of the manuscript.

Requests for reprints should be sent to Anthony J. Cuvo, Rehabilitation Institute, Southern Illinois University, Carbondale, Illinois 62901.

EXPERIMENT 1

Childs (1983) investigated the effect of the "cover write" method to teach spelling to children with mental retardation. This procedure involved initially presenting a written word as a model. The student then wrote the word several times, covered the model and all words written, wrote the word again without the model visible, and then self-evaluated the written response. The cover write method included a total of 10 written responses per trial for each word to be learned. Childs concluded that children with mental retardation learned to use the cover write method as a strategy; however, he did not evaluate the effects that the procedure had on actually learning to spell the words.

A similar procedure, termed the "cover copy compare" method, has been used to promote the acquisition or fluency of arithmetic (Skinner, Bamberg, Smith, & Powell, 1993; Skinner, Ford, & Yunker, 1991; Skinner, Shapiro, Turco, Cole, & Brown, 1992; Skinner, Turco, Beatty, & Rasavage, 1989), spelling (Murphy, Hern, Williams, & McLaughlin, 1990), and geography (Skinner, Belfiore, & Pierce, 1992) tasks with students who were mentally retarded, learning disabled, or behaviorally disordered. Social validation also indicated that students preferred the cover copy compare method to traditional methods of instruction (Murphy *et al.*, 1990). The add-a-word method, which has been used successfully to teach spelling, essentially is the same as the cover write and cover copy compare procedures (McLaughlin, Reiter, Mabee, & Byram, 1991).

The current study evaluated the role of the cover component as used in the cover write, cover copy compare, and add-a-word methods. The stimulus-covering procedure is an integral component of these instructional procedures. When learning spelling, the stimulus word initially is neutral and does not evoke the correct spelling response. Repeated prompt presentation, practice, and immediate feedback, as well as self-management procedures, may explain the possible effi-

cacy of instruction, and stimulus covering may not add to the effect of these procedures. The cover write method, therefore, was compared to a procedure that did not include a cover step, but was equivalent in the number of times the word was practiced (*i.e.*, write method).

In addition, an oral method was included that required students to practice spelling words orally an equal number of times that words were written in the cover write and write methods. It has been suggested that response topographies should be the same during training and assessment (Greenwood *et al.*, 1984). Research has shown that there was no difference between oral and written practice on a trials-to-criterion dependent variable, but oral practice was superior with respect to time to criterion (Van Houten & Van Houten, 1991). This research, however, did not provide a test of the Greenwood *et al.* recommendation, because participants were tested only in the modality in which they practiced. By comparing the oral method to the cover write and write methods in the present study, the effects of oral versus written practice could be determined when the assessment for all conditions was written.

Finally, in previous research involving the cover write or cover copy compare procedures, generalization testing typically was absent; therefore, the current study implemented a generalization test to determine whether acquisition of spelling would generalize differentially for the cover write, write, and oral methods.

METHOD

Participants

Four rehabilitation clients receiving independent living skills training in a university-affiliated rehabilitation facility participated. Participant characteristics are shown in Table 1.

Screening Procedures and Materials

The experimenter generated an initial pool of 279 words by reviewing independent living class materials at participants' training facility,

Table 1
Participant Characteristics

Subject	Age	IQ	Spelling grade equivalent
Experiment 1			
Kent	20	75 (WISC-R)	3.0 (WRAT)
Erin	20	80 (WISC-R)	3.7 (PIAT-R)
Wayne	19	85 (WISC-R)	2.8 (WRAT)
Tami	20	70 (WISC-R)	2.9 (WRAT)
Experiment 2			
Kitty	17	62 (WISC-R)	2.0 (Woodcock-Johnson)
Jack	17	47 (WISC-R)	1.6 (PIAT-R)
Kent	17	68 (Stanford-Binet)	10.5 (PIAT-R)
Lance	17	40 (WISC-R)	3.0 (PIAT-R)
Experiment 3			
Bobby	16	86 (WISC-R)	3.2 (WRMT-R)
Brent	15	90 (WISC-R)	5.0 (PIAT-R)
Martin	14	95 (WISC-R)	3.5 (PIAT-R)
Peter	17	105 (WISC-R)	8.0 (PIAT-R)
Jeff	15	104 (WISC-R)	7.0 (WRMT-R)
Experiment 4			
			Reading grade equivalent
Jennifer	15	40 (WISC-R)	0.7 (PIAT-R)
Don	14	65 (WISC-R)	1.0 (PIAT-R)
John	15	49 (WISC-R)	1.4 (PIAT-R)
Mike	15	55 (WISC-R)	1.4 (PIAT-R)
Beth	15	49 (WISC-R)	2.5 (PIAT-R)

WISC-R = Wechsler Intelligence Scale for Children—Revised.

WAIS-R = Wechsler Adult Intelligence Scale—Revised.

WRAT = Wide Range Achievement Test.

PIAT-R = Peabody Individual Achievement Test—Revised.

WRMT-R = Woodcock Reading Mastery Test—Revised.

speaking to the teachers of these classes, going to stores in the community, and using a list of community-referenced words generated in another city (J. Lalli, personal communication, June 1993). This pool of words was typed on two sheets of paper. To begin screening, the experimenter arbitrarily selected a word on the list, stated the word aloud, and asked participants to write it. When participants misspelled 60 words in writing, they were asked to spell the same words aloud. Screening continued until an individualized pool of 60 misspelled words was developed for each participant. In order to be included in the pool, words had to be misspelled on both written and oral tests on three independent test trials. From each participant's pool of words, 15 words were randomly

assigned to each of the three experimental conditions, and the remaining 15 served as no-training control words.

Participants were given sheets of green lined paper, a pencil with an eraser, and three laminated sheets of paper listing responses to be performed for each of the three experimental conditions. These latter sheets were used to prompt the responses for each experimental condition and to cover the written stimuli and responses in the cover write condition. Stimulus words were typed using a Macintosh Plus® computer (Geneva typeface, 18 point), laser printed, and cut into sheets (1 in. by 2 in.). A performance feedback chart was developed for each participant that showed the training words listed under each condition heading.

Table 2
Experiment 1 Procedures

Cover write method	Write method	Oral method
1. Look at word, say it	1. Look at word, say it	1. Say the word out loud
2. Print the word two times	2. Print the word three times	2. Read the letters of the word out loud 10 times
3. Cover, print one time	3. Check your work	
4. Check your work	4. Print the word three times	
5. Print the word two times	5. Check your work	
6. Cover, print one time	6. Print the word four times	
7. Check your work	7. Check your work	
8. Print the word three times		
9. Cover, print one time		
10. Check your work		

Setting and General Procedures

Participants were tested and trained individually, with the experimenter and participant sitting side by side at a desk in an office at the participants' service facility. Sessions were held four times per week for approximately 30 min each. The number of instructional sessions varied within conditions for each participant, because the acquisition criterion was number of trials to criterion. Every fourth session was audiotaped for calculating interscorer agreement.

At the beginning of the first session, the experimenter randomly chose four words, one from each of the three experimental conditions and one from the no-training control condition, for testing and training. Each subsequent session began with a written and oral assessment. The participant printed (written assessment) and spelled aloud (oral assessment) the words that were trained during the previous session. The performance feedback chart was presented if any of the words met the acquisition criterion. When a word met criterion, the participant crossed out the word on the feedback chart and received praise. The sequence of the written and oral assessments was counterbalanced across sessions.

After all four words had been assessed, training was initiated. The participant was informed of the instructional condition associated with each word when it was presented. When the first training trial was completed for all words,

two additional training trials were implemented in an identical manner each session. Praise was delivered contingent on the participant's general cooperation.

The acquisition criterion was correctly spelling a word on two consecutive written assessments, which could be achieved in a minimum of two sessions (i.e., only one assessment trial took place at the beginning of each session). When the acquisition criterion was attained, the word was removed from the training pool and was replaced by another word randomly assigned to the same experimental condition. At the end of the experiment, each participant was orally administered several social validation questions that inquired about various preferences for the experimental conditions.

Experimental Design

Three experimental conditions (i.e., cover write, write, and oral) were presented in an alternating treatments design. In addition, no-training control words were tested throughout the experiment. The sequence of the three training conditions was determined randomly, with the constraint that each condition must occur once during every three consecutive trials within a session.

Cover write method. Participants were trained initially to use the cover write method. The 10 steps of the procedure, shown in Table 2, were presented in written and oral form. Subsequent-

ly, each of the 10 steps was modeled and verbally described in succession using a practice word that the participant already knew how to spell.

The participant was then verbally instructed step by step to perform the cover write method for a different practice word. Cover-the-word steps (3, 6, and 9 in Table 2) involved placing the laminated paper that listed the response sequence over the model word as well as the words printed by the participant. For the check-work steps (4, 7, and 10 in Table 2) the trainer said, "See if the words you printed look like the one at the top of your paper. If they do, tell me. If they are not spelled right, draw a line through the word, and print it correctly next to the misspelled word." Error correction was praised.

Praise was provided at the end of this initial practice training trial. The participant's questions were answered, and a second practice trial occurred using a different word. The same procedures were followed until the participant responded correctly on all 10 steps without error on one practice trial. Praise was provided at the end of the 10 steps if all of them were performed correctly. After this initial training, the participant used this method to learn 15 experimental words.

Write method. As can be seen in Table 2, the write method was similar to the cover write method but did not include the covering procedure. The write method was equated to the cover write method with respect to the number of times the words were spelled and checked as well as all other relevant variables.

Oral method. Words were presented on the top of the lined paper as in the other conditions. Participants were asked to say the word and spell the letters of the word aloud 10 times. This condition was equated to the other two conditions with respect to the number of times the word was practiced, but the practice was oral and not written. Errors were corrected immediately, and praise was presented at the end of each trial.

Control words. Words in this condition had been misspelled three times during screening, both orally and in writing, but did not receive training. Each word in the control condition was randomly paired to a word in the oral condition as a means of determining how many trials it would be assessed. A new control word was introduced for testing when its counterpart in the oral condition met the acquisition criterion. One control word was assessed both in writing and orally immediately after the three words in the experimental conditions were assessed.

Generalization

Generalization was assessed in the context of spelling the training words in a sentence. The experimenter developed a list of simple sentences containing one or more of the training words presented in a community-referenced context (e.g., "I am *qualified* for the job"). The experimenter stated the sentence aloud and asked the participant to write it. Testing took place before and after training. Only the training word in each sentence was scored for accuracy of spelling.

Dependent Variables

The primary acquisition dependent variable was the total number of written assessment trials required to meet the acquisition criterion on all 15 words in each of the three experimental conditions. The generalization dependent variable was the percentage of training words spelled correctly in each experimental condition during generalization testing.

Interscorer Agreement

A secondary scorer, who was unfamiliar with the purpose of the experiment, scored 25% of the written and oral assessment trials. The experimenter and secondary scorer collaboratively scored one session from a pilot subject, then independently scored additional trials until 100% agreement occurred on three consecutive trials.

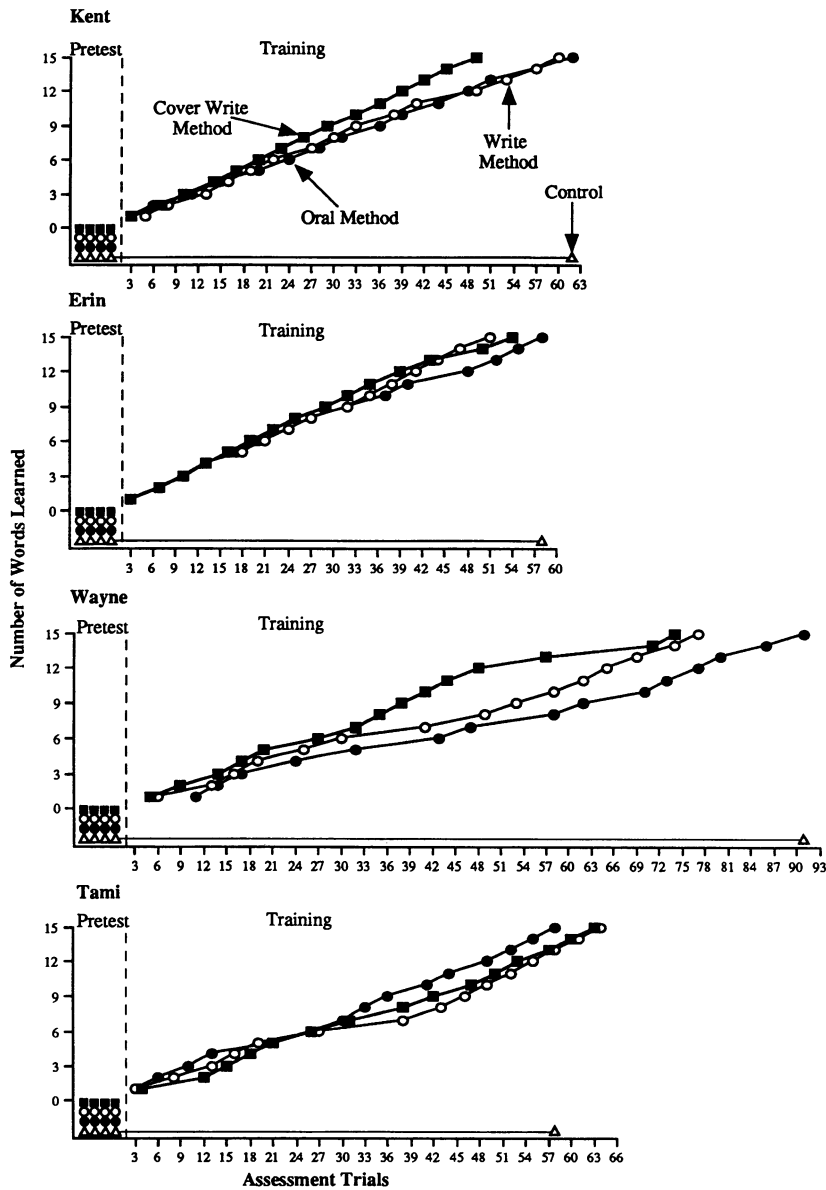


Figure 1. Number of assessment trials required to learn the 15 spelling words in each of the three experimental conditions by rehabilitation clients in Experiment 1.

For the cover write and write methods, participant data sheets were used to check scoring. For the oral method, the audiotape was used to determine interscorer agreement. An agreement was recorded when the experimenter and secondary scorer concurred that a word was spelled either correctly or incorrectly. Interscorer agreement was calculated as the number of agree-

ments divided by the number of agreements plus disagreements multiplied by 100%. Mean interscorer agreement was 97% for both written and oral assessments.

RESULTS

Figure 1 shows acquisition across assessment trials for the 4 participants. Each data point

represents the assessment trial on which the participant met the acquisition criterion for the word number indicated on the abscissa. The figure shows generally rapid acquisition for the three experimental conditions, with relatively small and variable differences among these conditions. In contrast to these training data, participants responded incorrectly to 100% of the control words that received no training. Acquisition, therefore, was a function of the experimental conditions.

The minimum of 45 assessment trials was needed to meet the acquisition criterion on all 15 words in each of the three experimental conditions. Kent required only 49 assessment trials in the cover write condition; however, he took 60 and 62 trials in the write and oral conditions, respectively. Wayne's overall acquisition rate was substantially slower than Kent's. He showed minimal difference between the cover write (73 trials) and write conditions (75 trials); the oral condition took longer (88 trials). Erin showed small differences among the three conditions. Acquisition was fastest in the write condition (51 trials) and was slightly slower in the cover write (54 trials) and oral (58 trials) conditions. Tami's acquisition was only marginally faster in the oral condition (58 trials) compared to the cover write (63 trials) and write (64 trials) conditions. Although Kent and Wayne performed somewhat better in the cover write condition than in the write condition, the data taken as a whole do not provide strong support for the enhanced effectiveness or efficiency of the cover component. Likewise, the results do not show that written practice is superior to oral practice when the assessment required a written response.

Generalization was tested when participants spelled the training words in sentences at the end of the experiment. Spelling accuracy ranged from 63.3% to 64.4% for all participants, compared to 0% correct on the pretest. There was little variability in generalization among the three experimental conditions; therefore, these data are not included in the figure.

When participants were interviewed at the conclusion of the experiment, 2 of the 4 indicated that they had no preference for any of the three experimental conditions. One participant responded that he preferred the cover write condition. All participants said that the oral condition was their least favorite and that the cover write condition helped them learn the words the best.

EXPERIMENT 2

Experiment 1 suggested that practice and reinforcement were the essential components of the cover write procedure. The covering component generally was not associated with enhanced learning over and above response practice. If learning occurs through practice, how much practice is necessary? Does more practice promote either more or faster acquisition?

Past research typically has examined the efficacy of only one frequency of practice, usually compared to a no-treatment baseline (e.g., Ollendick et al., 1980; Schumaker & Sherman, 1970). Few comparative studies have been conducted, and those on the duration of overcorrection have shown that this variable has not been consistently related to response reduction. Some results have shown that increasing the duration of overcorrection increases its effectiveness for reducing inappropriate behaviors (Foxy & Azrin, 1973; Ollendick & Matson, 1976; Sumner, Meuser, Hsu, & Morales, 1974), but other data are not supportive of that finding (Carey & Bucher, 1983; Conley & Wolery, 1980). A literature review did not reveal any studies that have evaluated the effects of differences in the number of times subjects were required to practice various tasks (Miltenberger & Fuqua, 1981).

Therefore, a parametric study was conducted to investigate the effectiveness of amount of practice in the context of the copy cover compare procedure on the spelling performance of students with autism and mental retardation. The copy cover compare method was used,

once again, because it is an established instructional technique with a supportive body of research, and Experiment 1 did not show a strong preference for either of the other two conditions.

METHOD

Participants

Three boys and 1 girl attending a special education class serving students with mild and moderate mental retardation at a local high school participated. Inclusion in the study was based on students having legible writing. Although all participants were in a self-contained classroom, they were integrated to varying degrees in regular education classes. Participant characteristics are described in Table 1.

Materials and Screening

Initial screening to obtain a pool of sight vocabulary words that participants could read correctly but not spell was conducted in a manner similar to that in Experiment 1. The words were derived from the *Children's Writer's Word Book* (Mogilner, 1992) that included lists of words at grade levels kindergarten to 6. A pool of 60 words was identified that participants misspelled on three independent test trials. The words were randomly assigned to four conditions, 15 in each of the three experimental conditions, and 15 in the no-training control condition. The words were equated as much as possible for number of letters and grade levels. A performance feedback chart was developed and implemented for each participant similar to that used in Experiment 1.

Participants were given a pencil with an eraser, sheets of white lined paper, and a different laminated piece of cardboard (the same size as the paper) for each training condition. The latter served as a prompt for the number of written practices required in a condition and as a cover for the words. One side of each cardboard listed instructional steps for its condition, typed using a Macintosh SE® computer (New Cen-

tury Schlbk typeface, size 14) and laser printed. The stimulus words were also typed using the same computer (Times typeface, size 48) and laser printed. The words were then cut (2.5 in. by 4 in.) and paperclipped to the top of the participant's paper on which the spelling was practiced.

Setting and Sessions

Training sessions were conducted in the school library, 4 or 5 days per week, for approximately 30 min per session. Each session consisted of an initial assessment of the words trained during the previous session, maintenance testing if necessary, and two subsequent training trials with a 3-min break between them. A training trial consisted of instruction on one word in each of the three experimental conditions. Participants were screened, tested, and trained individually, with the participant and experimenter sitting side by side at a table. Every fourth session was used to check inter-scorer agreement.

General Procedures

Initially, participants were trained to perform the CCC10X procedure shown in Table 3. Subsequently, the experiment proper began. At the beginning of each training session, participants were tested on words trained during the previous session. The experimenter dictated words, and the participant printed them on white lined paper. If a word was spelled correctly for two consecutive test trials, the acquisition criterion was met and the word was removed from training. The general procedures, acquisition criterion, and use of the performance feedback chart were the same as in Experiment 1. At the end of each session, participants were given social, edible, and tangible rewards contingent on their cooperation.

Experimental Design and Conditions

An alternating treatments design was implemented to determine the effects of the three

Table 3
Experiment 2 Procedures

CCC5X	CCC10X	CCC15X
1. Look at the word and say it	1. Look at the word and say it	1. Look at the word and say it
2. Copy word two times	2. Copy word two times	2. Copy word two times
3. Cover and write one time	3. Cover and write one time	3. Cover and write one time
4. Check work	4. Check work	4. Check work
5. Copy word two times	5. Copy word two times	5. Copy word two times
6. Check work	6. Cover and write one time	6. Cover and write one time
	7. Check work	7. Check work
	8. Copy three times	8. Copy three times
	9. Cover and write one time	9. Cover and write one time
	10. Check work	10. Check work
		11. Copy word two times
		12. Cover and write one time
		13. Check work
		14. Copy word two times
		15. Check work

experimental conditions and the no-training control condition. The order of presenting experimental conditions was randomized on each training trial as described in Experiment 1. The three experimental conditions were identical except for the number of times participants practiced the stimulus words to be learned: cover copy compare five times (CCC5X), cover copy compare 10 times (CCC10X), and cover copy compare 15 times (CCC15X). Otherwise, conditions were identical to the cover write condition in Experiment 1.

Participants were presented with the words and performed the procedures shown in Table 3. In the three experimental conditions, participants practiced spelling five, 10, or 15 times, respectively. In the no-training control condition, words received no training and were tested at the conclusion of the experiment. The acquisition dependent variable was similar to that in Experiment 1.

Interscorer Agreement

Interscorer training and scoring procedures were similar to those in Experiment 1. The percentage agreement was 100% across all participants and words in all experimental conditions.

RESULTS

Figure 2 shows acquisition data as a function of assessment trials for the 4 participants. Participants rapidly learned to spell all the target words with minimal differences among conditions. Learning in the three experimental conditions was far superior to that in the no-training control condition, in which only one word was learned by each of 2 participants, and no words were learned by the other 2 participants. However, there was some variability in the rate of learning across participants. Kent learned the fastest and Kitty was the slowest.

Kent and Kitty maintained all words learned in each condition during the three weekly follow-up tests. Lance maintained 44 of the 45 words (98%); he relearned one word from CCC5X within two training sessions. Jack maintained 41 of the 45 trained words (91%). With additional training, Jack rapidly reacquired the four words that had not been maintained (one from CCC5X, one from CCC10X, and two from CCC15X).

EXPERIMENTS 3 AND 4

Experiment 2 showed that students with developmental disabilities acquired spelling behav-

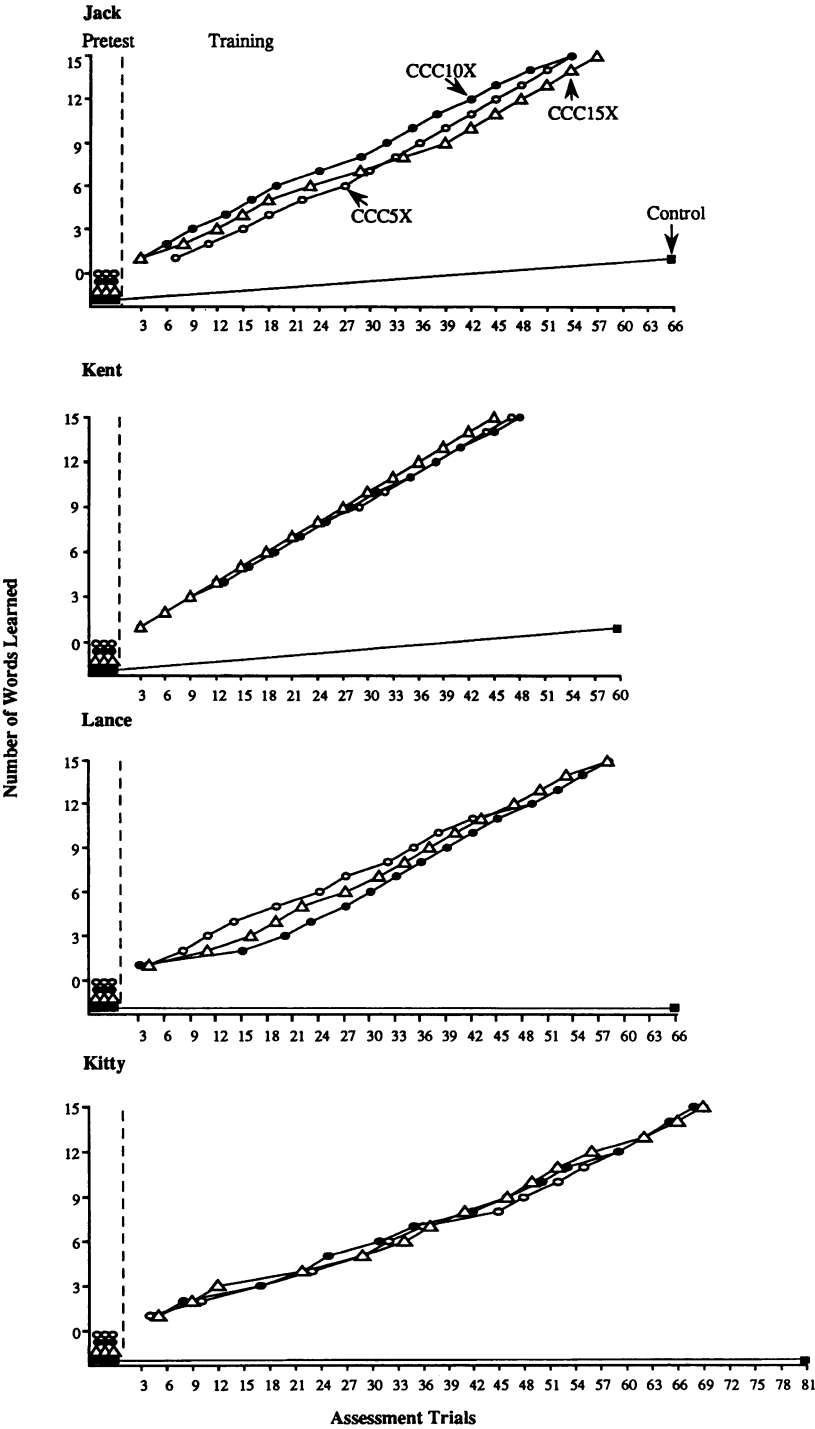


Figure 2. Number of assessment trials required to learn the 15 spelling words in each of the three experimental conditions by students with developmental disabilities in Experiment 2.

ior quite rapidly regardless of amount of practice (i.e., five, 10, 15 times) and maintained those responses over several weeks. One purpose of Experiments 3 and 4 was to test whether instruction could be even more efficient by using a one-time practice condition. An additional purpose was to investigate the effect of relevant and irrelevant response practice. A defining characteristic of overcorrection is that the repetitive practice is topographically similar to the behavior that is being modified (Foxx & Bechtel, 1982). The logic of that requirement is to support the educative function of the procedure. If learners acquire a target behavior when their practice is irrelevant to that behavior, then it could be argued that acquisition is affected by variables other than positive reinforcement. Perhaps, the target behavior is strengthened by negative reinforcement in the context of avoidance learning.

Two studies (Axelrod et al., 1984; Rodgers & Iwata, 1991) have investigated the comparative effects of task-relevant and irrelevant practice. Axelrod et al. examined whether it was necessary to use topographically similar overcorrection (i.e., write the correct spelling of each erroneous response on the pretest, identify the phonetic spelling and part of speech, locate the dictionary definition, and use the word correctly five times) or whether topographically dissimilar overcorrection (i.e., practice as described above for words not on the pretest) would produce improvement in spelling performance of 1 mentally retarded and 1 emotionally disturbed student. Results indicated that both the relevant and irrelevant practice conditions produced significant and approximately equal improvements in spelling when compared to baseline levels.

Rodgers and Iwata (1991) attempted to answer the same conceptual question in the context of a match-to-sample discrimination training task for adults with severe and profound mental retardation. Three conditions with different error contingencies were administered using an alternating treatments design. In the differential reinforcement condition (baseline), no

consequences were delivered following an incorrect response. When errors occurred in the practice condition, participants continued to respond in a topographically relevant manner to experimental stimuli until one correct response occurred. For the final "avoidance" condition, additional training trials were conducted on task-irrelevant stimuli following an incorrect response. Results showed that all participants made noticeable progress in baseline as well as in the two training conditions. There was relatively small separation in the data curves for experimental conditions, which resulted only after a large number of training trials. The investigators recommended additional research to investigate the function of error-correction procedures on various training tasks.

The purpose of Experiments 3 and 4, therefore, was to extend the research on task relevance of response practice with students with behavior disorders and mental retardation. Would students learn to spell or recognize sight words when practice following an error was irrelevant to the word to be learned as well as when practice was relevant? Three practice procedures and a no-treatment control condition were compared for promoting the acquisition and maintenance of spelling by students with behavior disorders (Experiment 3) and sight words by students with mental retardation (Experiment 4). A one relevant practice condition (i.e., one-time practice of the target word) was implemented to measure the effects of minimal practice on responding, and a five relevant practice condition determined whether additional task-relevant practice (i.e., five times) would be even more effective as a result of additional positive reinforcement of the target response or negative reinforcement from avoidance of the practice requirement. A third condition, five irrelevant practice, was included to examine whether it was necessary for practice to be relevant to the target words. Because this condition did not provide positive reinforcement of the target responses, these words could be

learned by negative reinforcement from avoidance of irrelevant practice.

METHOD

Participants

Participants in Experiment 3 were 5 male adolescents attending a self-contained class for students with behavior disorders at the same high school that Experiment 2 participants attended. The majority of these participants were mainstreamed in regular education classes throughout the school day. Participants in Experiment 4 were 2 females and 3 males from the same classroom as those in Experiment 2. Participants who had verbal skills and were at the lowest reading grade levels were nominated by the teacher. Participants were integrated in regular education classes on an individually determined basis. Participant characteristics for both experiments are described in Table 1.

Settings and Sessions

Sessions took place either in the school library four to five times per week, approximately 20 min per session (Experiment 3), or in the cafeteria four times per week, approximately 10 min per session (Experiment 4). Participants were screened, tested, and trained individually. Each session consisted of an initial assessment of the words to be trained that day, followed by three training trials, and finally a second assessment of the words just trained. All sessions were audiotaped, and every fourth session was used for calculating interscorer agreement.

Materials and Screening

Participants for Experiment 3 were screened to attain a pool of sight vocabulary words that could be named but not spelled correctly as in previous experiments. The words were derived from the teacher's edition of *Spelling Skill* (Fletcher, Endres, Urban, Hilbert, & Grant, 1983). Screening for Bobby, Brent, Peter, and Jeff was conducted with words at their tested reading grade equivalents. Martin's words were

obtained from the same text, except they were four grade levels above his current reading level. The selection of more difficult words was undertaken when the results for the previous participants suggested that there were minimal acquisition differences among experimental conditions. Screening yielded a total of 60 individually determined words for each participant; 45 of the words were randomly selected to serve as training words, 15 in each of the three experimental conditions, and the remaining 15 words were used as no-training controls.

An additional 15 words were selected as no-training words and were paired to the 15 training words in the five irrelevant practice condition. For 4 students, the 15 no-training words were obtained from a word list two grade levels above their present spelling levels. The no-training words had more letters (i.e., 10) than the training words to increase their possible difficulty (aversiveness) and avoidance. Martin's no-training words also were four grades higher. For all participants, the training words in each of the three experimental conditions were equated according to number of letters in each word.

Participants for Experiment 4 were screened individually as described above to obtain a pool of words that they could not read at their tested reading grade levels. For Jennifer, Don, John, and Mike, words were derived from a curriculum guide, *Success in Spelling: A Course in Spelling for Grades 1-6* (1982), that included numerous lists of words at each grade level up to Grade 6. Words were chosen that matched their reading grade equivalents as closely as possible. Beth's words were chosen from *The Living World Vocabulary* (Date & O'Rourke, 1981) four grades above her tested reading grade equivalent to increase their difficulty.

From each participant's pool of words, 45 were randomly chosen for training, and the remaining 15 words served as no-training control words. The 60 words were equated according to number of letters in each word. An additional 15 words were selected from the same sources to be used as no-training words in the

five irrelevant practice condition. These 15 words, each paired with a target word, had more letters than the training words to increase their presumed difficulty. Words were printed using Microsoft® Word software on a Macintosh® computer (Times typeface, 60 point) and were glued to index cards (3 in. by 5 in.). The name of the experimental condition to which the word was assigned was written on the back of the index card. The visual feedback chart was used in the same manner as described in previous experiments.

Procedure

The procedures for the two experiments were similar, with the exception that in Experiment 3 the words were presented to teach spelling to students with behavior disorders, and in Experiment 4 the words were presented to teach sight vocabulary to students with mental retardation. The experimenter randomly chose six words, two from each of the three experimental conditions, to begin assessment and training. Throughout training, two words from each of the three conditions were tested and instructed each session until all 15 words in each condition were learned.

The participant was initially tested on the eight words (six training and two no-training) to be used that session. For spelling instruction, words were presented orally; for sight word instruction, words were presented on index cards. During assessment, neither prompts nor performance feedback was provided; however, praise was delivered on an arbitrary schedule for the participant's cooperation.

After all eight words had been assessed, training was initiated. When a word was presented, the participant was informed of the consequences for making an error based on the experimental condition to which the word had been assigned. If the participant did not respond within 5 s or made an incorrect response, the correction procedure was applied, and the next word was presented. When all six words were instructed in this manner and the first

training trial concluded, a second and then a third training trial was implemented in an identical fashion. At the end of the third training trial, the participant was assessed again without prompts or consequences. The six training words plus any no-training words used in the five irrelevant practice condition were presented once, in random order, for assessment.

The acquisition criterion for a word was correct responses during three consecutive assessment trials. The minimum number of sessions to achieve this criterion was two (i.e., there were two assessment trials each session, one prior to and one following the three training trials). The participant could err on all previous training trials, in principle, but if the word was correct on three consecutive assessment trials, the word met the acquisition criterion. When the acquisition criterion was met for a word, it was replaced by another word randomly assigned to the same experimental condition. The participant was praised, was told that he or she would have to practice the word only one time per week, and was told to cross off the word on the visual feedback chart. Training and assessment continued until all 15 words in each condition were learned.

A maintenance check was conducted at the end of the last session each week, immediately after the second assessment trial. All of the words that had previously met the acquisition criterion were presented for testing. No prompts or performance feedback was provided during maintenance checks. The maintenance criterion was correct responding during four consecutive maintenance probes (i.e., 4 weeks). If an error occurred once during a maintenance probe, the word was reintroduced for training in its original experimental condition after all other words in that condition had been trained. No-training words used for the five irrelevant practice condition were also probed for maintenance, but they were not retrained when errors occurred.

At the end of the experiment, each participant (except Beth, who was unavailable) was

orally administered several social validation questions that inquired about various preferences for the experimental conditions.

Experimental Design and Conditions

For both experiments, an alternating treatments design was used to determine the effect of the three practice procedures (one relevant practice, five relevant practice, and five irrelevant practice) on the number of spelling or sight vocabulary words learned. In addition, there was a no-treatment control condition in which 15 untrained words were tested at the conclusion of the experiment. Training trials consisted of the words from each of these experimental conditions presented randomly, as described above. Praise was used as a consequence in all three experimental conditions for correct responses.

One relevant practice. Prior to a word being presented in this condition, the participant was told, "If you miss this word, you will have to spell or say the correct word one time." Incorrect spelling responses or no response after 5 s was followed by the experimenter saying, "No, that is incorrect. Let me give you the correct spelling." The experimenter then modeled the correct written response, and the participant wrote the word once. If a participant still responded incorrectly, the correct response was written again, and the participant was required to emit one correct written response before receiving praise. For sight vocabulary errors, the practice procedure was similar. The experimenter corrected the error, and the participant verbally stated the word correctly once.

Five relevant practice. Before presentation of the first word, the participant was told, "If you miss this word, you will have to practice the right word five times." Error-correction procedures were similar to those described above, except that the participant was required to emit the correct written (spelling) or oral (sight word) response five times.

Five irrelevant practice. Before presentation of the first word, the participant was told, "If you

miss this word you will have to practice another word five times that is much harder and will not help you finish these words any faster." When an incorrect response was made, the experimenter provided the correct response, but the participant did not practice it. The same general error-correction procedure was used as in the five relevant practice condition, with one exception; a no-training word was used for practice instead of the training word. When the acquisition criterion for the training word was met, both it and its paired no-training word were discontinued from assessment and training.

Dependent variable. The acquisition dependent variable was similar to those used in previous experiments.

Interscorer Agreement

Agreement was calculated as in previous experiments. For Experiment 3, interscorer agreement was 100%. For Experiment 4, the mean was 96% (range, 92% to 100%); no procedural variations from the established protocols were observed on an independent variable check.

RESULTS

Experiment 3

The number of assessment trials required to meet the acquisition criterion on all 15 words for each experimental condition is shown in Figure 3. The minimum number of trials possible was 32. Peter and Bobby required only 32 or 33 trials in each condition, and Jeff and Brent took 34 to 36 trials per condition; Martin took 36 to 42 trials per condition and required only 4 days of additional training to learn all 45 of his more difficult words (i.e., four grade levels higher). All participants responded incorrectly to the 15 no-training words when post-tested. These results show small differences among experimental conditions. One-time relevant practice generally was as effective as either five-times relevant or irrelevant practice.

A secondary dependent variable was the per-

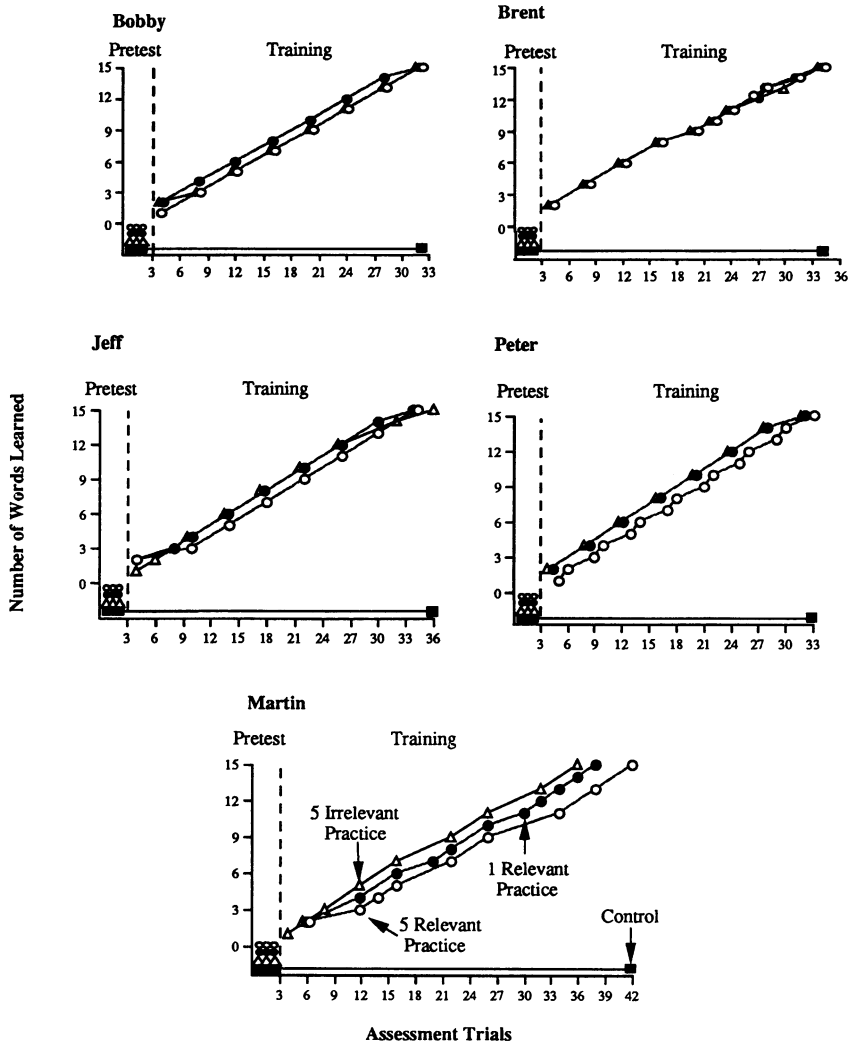


Figure 3. Number of assessment trials required to learn the 15 spelling words in each of the three experimental conditions by students with behavior disorders in Experiment 3.

centage of no-training words in the five irrelevant practice condition spelled correctly per assessment trial. The percentage, which ranged from 8.8% to 44.1% across participants, was significantly lower than for training words acquired in the three experimental conditions. Although participants practiced these task-irrelevant words and did not practice the target words, accuracy was greater for the latter. Practice of the task-irrelevant words in the absence of contingencies to learn did not result in more words spelled correctly.

Two dependent variables pertinent to response maintenance were the percentage of training words spelled correctly per condition and the percentage of no-training words in the five irrelevant practice condition spelled correctly on four consecutive weekly maintenance tests. Peter, Jeff, Bobby, and Brent all maintained the majority (i.e., 87% to 100%) of words on all four maintenance assessments across the three experimental conditions without additional training.

Martin was the only participant who showed

maintenance variability across the three experimental conditions. He maintained 100% of the words in the five relevant practice condition for the first 3 weeks of assessments and 93% during the fourth assessment. The results in this condition are comparatively better than or equal to those of the other participants who had easier words. For the other two experimental conditions (i.e., one relevant and five irrelevant practice) Martin's response maintenance dropped to about two thirds of the words spelled correctly during the third and fourth assessments. All words were maintained for an additional 4 weeks after retraining. The percentage of no-training words maintained in the five irrelevant practice condition ranged from 13.3% to 46.6% across participants, considerably lower than the training words.

Participants were interviewed at the conclusion of the experiment concerning their preferences for the three experimental conditions. Three of the participants responded that they preferred the one relevant practice condition; 2 indicated no preference. When asked which condition was their least favorite, all 5 participants said the five irrelevant practice condition. For the condition participants believed helped them learn to spell more words, 4 of the 5 responded, "five relevant practice"; 1 participant said he believed he learned to spell words the same regardless of experimental condition.

Experiment 4

Figure 4 shows that participants required only a few trials above the minimum of 32 to learn all the words. There was little difference among experimental conditions, except that Don learned the words in the five irrelevant practice condition slightly slower (required two more sessions) than the words in the other two conditions. Because of the termination of the school year, Beth was trained on 10 instead of 15 words per condition. Although her words were four grade levels above her current grade level, her acquisition was also rapid. All participants responded incorrectly to 100% of the

control words when posttested. Once again, increased frequency and relevance of practice were not any more effective than either lesser amounts or irrelevant practice.

The percentage of no-training words stated correctly in the five irrelevant practice condition ranged from 40.0% to 72.7%, which was considerably lower than for the training words that were not practiced. As in Experiment 3, practice without contingencies produced much more modest acquisition, and more words that were not practiced were acquired.

Only 2 participants failed to state correctly 100% of the words trained during the four weekly maintenance tests. Jennifer erred on one of her training words in the one relevant practice condition during Week 1; Don erred on one training word each in the one relevant and five irrelevant practice conditions during Week 1. After retraining, all words were maintained by Jennifer and Don. Beth maintained all 10 words during the first 2 weeks, but the third and fourth tests could not be completed because the school year ended.

Maintenance of the no-training words from the five irrelevant practice condition was substantially lower than that for the training words. Jennifer did not maintain any words; the other participants showed a progressive decline in maintenance across the 4 weeks, culminating in 27% to 67% correct responses.

Social validation results indicated that 2 participants said that the most helpful condition was five irrelevant practice, and 2 said five relevant practice. As in Experiment 3, all 4 participants reported that the five irrelevant practice condition was their least favorite. The one and five relevant practice conditions were each favored by 2 participants.

GENERAL DISCUSSION

The present series of experiments investigated variables that are relevant to response practice in the context of academic instruction for individuals with intellectual disabilities or be-

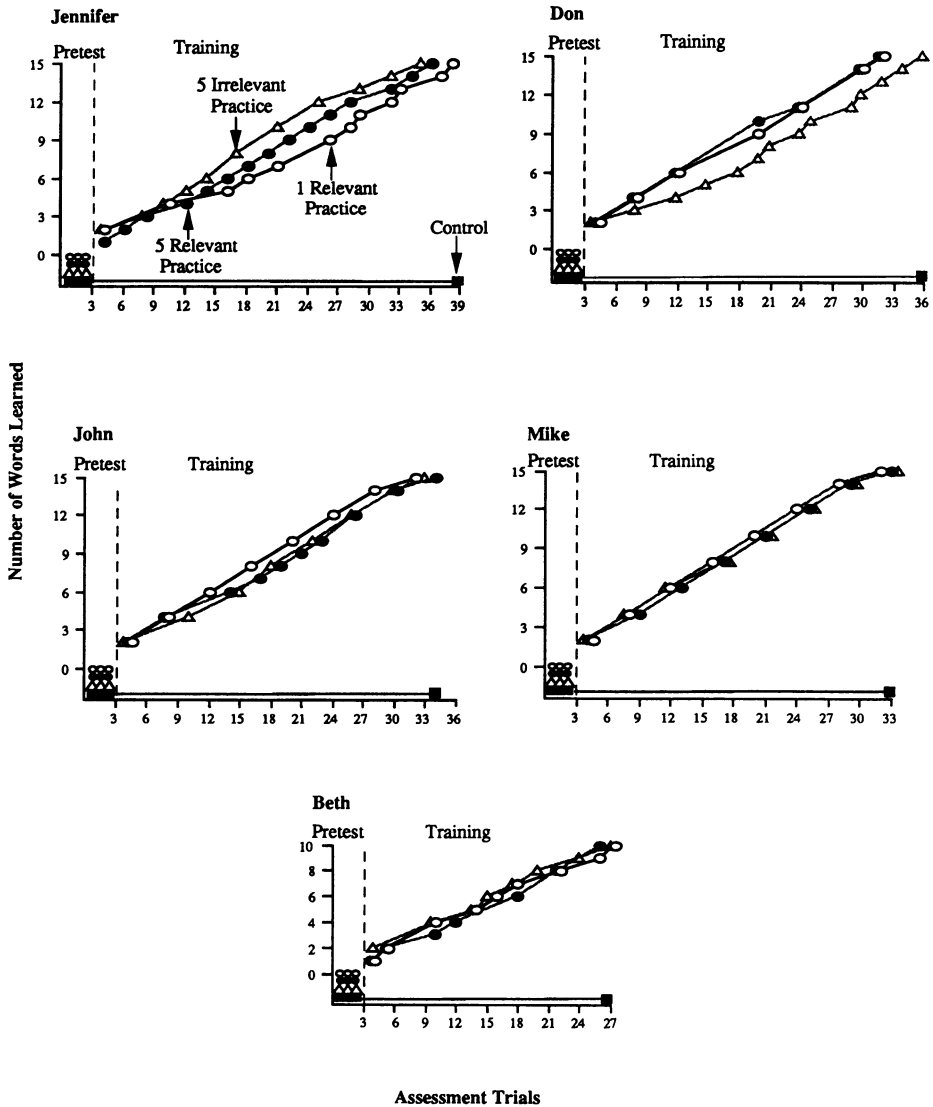


Figure 4. Number of assessment trials required to learn 15 sight vocabulary words in each of the three experimental conditions by students with mental retardation in Experiment 4.

havior disorders. Experiments 1 and 2 showed that the cover write and cover copy compare procedures, which are procedurally equivalent, resulted in rapid learning of spelling words by adult rehabilitation clients and adolescents with mental retardation. The component analysis in Experiment 1, however, indicated that with the possible exception of Kent, the cover procedure generally did not facilitate learning over and above that produced by response practice, which included response repetition, immediate

feedback, and self-correction. Also, oral and written practice were equally effective when assessment was written, with negligible differences in efficiency between the two types of practice.

Experiment 2 showed that there was no positive association between the frequency of response practice and measures of learning. This was the first known parametric analysis of the effect of frequency of response practice on learning academic tasks in the context of the

cover copy compare procedure. Participants learned just as well when practice was limited to five repetitions as with 10 or 15 repetitions; however, learning did not take place in the control condition in the absence of formal practice.

Experiments 3 and 4 showed that sufficient stimulus control obtained with one-time relevant practice. The additional practice in the five relevant practice condition, with its ostensibly greater positive and possibly negative reinforcement, was unnecessary for learning. This finding held for target words at the participants' grade levels, as well as for 2 participants when the words were four grade levels higher.

Interestingly, participants learned just as well when the experimenter corrected the error and participants engaged in five irrelevant practices. The lack of difference between relevant and irrelevant practice confirms the results of Axelrod *et al.* (1984) and, substantially, Rodgers and Iwata (1991). The lack of difference between irrelevant and relevant practice is a counterintuitive finding because irrelevant practice *per se* could not have served a direct educative function.

On the one hand, acquisition in the five irrelevant practice condition could be attributable to error correction by the experimenter and possibly to covert practice by the participant; however, mitigating against that possibility was the fact that the required overt irrelevant practice was incompatible with covert relevant practice. Alternatively, acquisition seemed to be enhanced by negative reinforcement of avoiding the repetitive task-irrelevant practice. All 10 participants said that the five irrelevant practice condition was their least favorite, subjectively confirming the possible aversiveness of the task.

The significance of the present research is that practice and reinforcement without further embellishment are powerful stimulus control procedures to promote the learning of spelling and sight vocabulary. Acquisition is not necessarily enhanced, however, either by the cover procedure or by additional repetition. Also, the results do not support the recommendation that

practice and assessment should be topographically similar (Greenwood *et al.*, 1984). Although task-irrelevant practice was just as effective as task-relevant practice, the social acceptability of the latter would be greater in most educational settings. Taken together, the results suggest the efficacy and procedural efficiency of moderate practice and positive reinforcement, even when the words to be learned are at an advanced level (as they were for 2 participants). Fortunately for teachers and students, more parsimonious instructional procedures may be as effective and efficient as more elaborate ones.

REFERENCES

- Axelrod, S., Kramer, A., Appleton, E., Rockett, T., & Hamlet, C. C. (1984). An analysis of the relevance of topographical similarity on positive practice of spelling errors. *Child and Family Behavior Therapy*, 6, 19-31.
- Baker, J. D. (1992). Correcting the oral reading errors of a beginning reader. *Journal of Behavioral Education*, 2, 337-343.
- Carey, R. G., & Bucher, B. (1983). Positive practice overcorrection: The effects of duration of positive practice on acquisition and response reduction. *Journal of Applied Behavior Analysis*, 16, 101-109.
- Childs, R. E. (1983). Teaching rehearsal strategies for spelling to mentally retarded children. *Education and Training of the Mentally Retarded*, 50, 318-320.
- Conley, O. S., & Wolery, M. R. (1980). Treatment by overcorrection of self-injurious eye-gouging in preschool blind children. *Journal of Behavior Therapy and Experimental Psychiatry*, 11, 121-125.
- Dalrymple, A. J., & Feldman, M. A. (1992). Effects of reinforced directed rehearsal on expressive sign language learning by persons with mental retardation. *Journal of Behavioral Education*, 2, 1-16.
- Date, E., & O'Rourke, J. (1981). *The living world vocabulary—a national vocabulary inventory*. Chicago: World Book-Childcraft.
- Fletcher, B., Endres, M. A., Urban, G. A., Hilbert, M. K., & Grant, M. (1983). *Spelling skill* (teacher's ed.). New York: Sadlier/Oxford.
- Fox, R. M., & Azrin, A. H. (1973). The elimination of self-stimulatory behavior by overcorrection. *Journal of Applied Behavior Analysis*, 6, 1-14.
- Fox, R. M., & Bechtel, D. R. (1982). Overcorrection. In M. Hersen, R. M. Eisler, & P. M. Miller (Eds.), *Progress in behavior modification* (Vol. 13, pp. 227-288). New York: Academic Press.
- Fox, R. M., & Jones, J. R. (1978). A remediation program for increasing the spelling achievement of ele-

- mentary and junior high school students. *Behavior Modification*, 2, 211–230.
- Greenwood, C., Dinwiddie, G., Terry, B., Wade, S., Stanley, O., Thibideau, S., & Delquadri, J. (1984). Teacher versus peer mediated instruction: An analysis of achievement outcomes. *Journal of Applied Behavior Analysis*, 17, 521–538.
- Heron, T. E., Okyere, B. A., & Miller, A. D. (1991). A taxonomy of approaches to teach spelling. *Journal of Behavioral Education*, 1, 117–130.
- Mabee, S. W. (1988). The effects of academic positive practice on cursive letter writing. *Education and Treatment of Children*, 11, 143–148.
- Matson, J. L., Esveltd-Dawson, K., & Kazdin, A. E. (1982). Treatment of spelling deficits in mentally retarded children. *Mental Retardation*, 20, 76–81.
- McLaughlin, T. F., Reiter, S. M., Mabee, W. S., & Byram, B. J. (1991). An analysis and replication of the add-a-word spelling program with mildly handicapped middle school students. *Journal of Behavioral Education*, 1, 413–426.
- Miltenberger, R. G., & Fuqua, R. W. (1981). Overcorrection: A review and critical analysis. *The Behavior Analyst*, 4, 123–141.
- Mogilner, A. (1992). *Children's writer's word book* (1st ed.). Cincinnati, OH: Writer's Digest Book, F & W Publication Inc.
- Murphy, J. F., Hern, C. L., Williams, R. L., & McLaughlin, T. F. (1990). The effects of the copy, cover, compare approach in increasing spelling accuracy with learning disabled students. *Contemporary Educational Psychology*, 15, 379–386.
- Ollendick, T. H., & Matson, J. L. (1976). An initial investigation into the parameters of overcorrection. *Psychology Report*, 39, 1139–1142.
- Ollendick, T. H., Matson, J. L., Esveltd-Dawson, K., & Shapiro, E. S. (1980). Increasing spelling achievement: An analysis of treatment procedures utilizing an alternating treatments design. *Journal of Applied Behavior Analysis*, 13, 645–654.
- Rodgers, T. A., & Iwata, B. A. (1991). An analysis of error-correction procedures during discrimination training. *Journal of Applied Behavior Analysis*, 24, 775–781.
- Schumaker, J., & Sherman, J. A. (1970). Training generative verb usage by imitation and reinforcement procedures. *Journal of Applied Behavior Analysis*, 3, 273–287.
- Singh, N. N. (1987). Overcorrection of oral reading errors: A comparison of individual and group-training formats. *Behavior Modification*, 11, 165–181.
- Singh, N. N., Farquhar, S., & Hewett, A. E. (1991). Enhancing the spelling performance of learning disabled students. *Behavior Modification*, 15, 271–282.
- Singh, N. N., & Singh, J. (1986). Increasing oral reading proficiency: A comparative analysis of drill and positive practice overcorrection procedures. *Behavior Modification*, 10, 115–130.
- Singh, N. N., & Singh, J. (1988). Increasing oral reading through overcorrection and phonic analysis. *American Journal on Mental Retardation*, 93, 312–319.
- Singh, N. N., Singh, J., & Winton, A. S. W. (1984). Positive practice overcorrection of oral reading errors. *Behavior Modification*, 11, 23–37.
- Skinner, C. H., Bamberg, H. W., Smith, E. S., & Powell, S. S. (1993). Cognitive, cover, copy, and compare: Subvocal responding to increase rates of accurate division responding. *Remedial and Special Education*, 14, 49–56.
- Skinner, C. H., Belfiore, P. J., & Pierce, N. (1992). Cover, copy, and compare: Increasing geography accuracy in students with behavior disorders. *School Psychology Review*, 21, 73–81.
- Skinner, C. H., Ford, J. M., & Yunker, B. D. (1991). A comparison of instructional response requirements on the multiplication performance of behaviorally disordered students. *Behavioral Disorders*, 17, 56–65.
- Skinner, C. H., Shapiro, E. S., Turco, T. L., Cole, C. L., & Brown, D. K. (1992). A comparison of self and peer-delivered immediate corrective feedback on multiplication performance. *Journal of School Psychology*, 30, 101–116.
- Skinner, C. H., Turco, T. L., Beatty, K. L., & Rasavage, C. (1989). Cover, copy, and compare: A method for increasing multiplication performance. *School Psychology Review*, 18, 412–420.
- Stewart, C. A., & Singh, N. N. (1986). Overcorrection of spelling deficits in moderately mentally retarded children. *Behavior Modification*, 10, 355–365.
- Success in spelling: A course in spelling for grades 1–6* (2nd ed.). (1982). Rochester, MN: Rochester Public Schools.
- Sumner, J. C., Meuser, S. T., Hsu, L., & Morales, R. G. (1974). Overcorrection treatment for radical reduction of aggressive-disruptive behavior in institutionalized mental patients. *Psychological Report*, 35, 655–662.
- Van Houten, R., & Van Houten, J. (1991). The effects of breaking new spelling words into small segments on the spelling performance of students with learning disabilities. *Journal of Behavioral Education*, 1, 399–411.

Received September 2, 1994

Initial editorial decision November 3, 1994

Revisions received February 7, 1995; February 24, 1995

Final acceptance March 15, 1995

Action Editor, Brian Martens